CULTURAL DIFFERENCES IN PREFERENCE OF AUDITORY EMOTICONS: USA AND SOUTH KOREA

Myounghoon Jeon

Michigan Technological University
Mind Music Machine Laboratory
Houghton, MI 49931, USA
mjeon@mtu.edu

Ju-Hwan Lee

Korean German Institute of Technology,
Interaction Laboratory,
Gangseogu, Seoul 157-030, Korea
jhlee@kgit.ac.kr

Jason Sterkenburg

Michigan Technological University
Mind Music Machine Laboratory
Houghton, MI 49931, USA
jtsterke@mtu.edu

Christopher Plummer

Michigan Technological University
Department of Visual and Performing Arts
Houghton, MI 49931, USA
cplummer@mtu.edu

ABSTRACT

For the last two decades, research on auditory displays and sonification has continuously increased. However, most research has focused on cognitive and functional mapping rather than emotional mapping. Moreover, there has not been much research on cultural differences on auditory displays. The present study compared user preference of auditory emoticons in two countries: USA and South Korea. Seventy students evaluated 112 auditory icons and 115 earcons regarding 30 emotional adjectives. Results indicated that they showed similar preference in the same category (auditory icons or earcons), but they showed different patterns when they were asked to select the best sound between the two categorical sounds. Implications for cultural differences in preference and directions for future design and research of auditory emoticons are discussed.

1. INTRODUCTION

For the last two decades, research on auditory display and sonification, the use non-speech sounds [1], has proliferated. Auditory icons [2] (representative part of sounds of objects or functions) and earcons [3] (ear + icons, short musical motives as symbolic representations of objects or functions) have been successfully applied to electronic devices as representative non-speech auditory feedback for user activity [e.g., 4, 5]. As tweaked speech cues, spearcons [6] (compressed speech) and spindex [7] (speech + index) have also shown enhanced performance and reduced workload with auditory menu navigation tasks in diverse contexts, such as desktop, mobile, and automotive environments. However, all these sonification approaches are based on cognitive and functional mappings as general HCI heavily depends on cognitivism. Given that sound and music are deeply related to human emotions, recent approaches try to embrace emotions and affect in the design of auditory displays. For example, musicons [16] (music + earcons) and lyricons [17] (lyrics + earcons) have been introduced to improve aesthetic aspects of the non-speech sound cues. A few studies have attempted to treat with emotional aspects of auditory icons [8] or earcons [9], but few studies compared affective effects of both auditory cues in a single study [exception, 10]. Another research gap is that an emotion study has relied simply on the valence dimension [positive – negative, e.g., 11]. Moreover, little research has focused on cultural differences on users' perception on auditory displays. To tackle these issues, the present paper assesses auditory emoticons composed of both auditory icons and earcons across two different countries, USA and South Korea. This systematic comparison will provide acoustic parameters of the emotional sounds, which will guide future design and implementation of auditory emoticons in user interfaces.

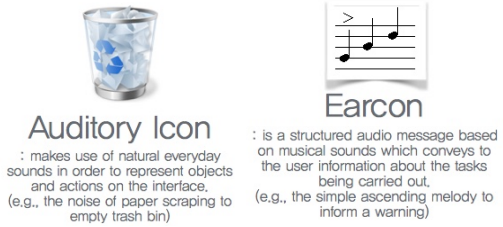

Figure 1: Examples of auditory icon and earcon.

2. DESIGN OF THE SOUNDS FOR 30 EMOTIONS

Sixteen college students, who major (or minor) in sound design or audio technology at Michigan Technological University (MTU), created in total 640 auditory icons and earcons for 30 affective adjectives (angry, boring, calm, cold, comfortable, confused, dark, delicate, depressed (sad), disgusting, dreamy, dynamic, fancy, free, fresh, harsh, impressive, intimate, lively, magnificent, modern, plain, pleasant (happy), scared (fearful), simple, soft, strong, surprising, uneasy, and warm) based on multi-phase design

panel discussions [12] under the two sound design experts' supervision. Sound-specific affective adjectives were selected from previous research using the statistical reduction processes (factor analysis and multi-dimensional scaling) [13, 14] and a couple of adjectives (disgusting and fearful) were added, to include basic six emotions [15] in the study. After completing iterative design panel sessions (3 times) and removing acoustically similar sounds, we selected 112 auditory icons and 115 earcons for further evaluations.

3. METHOD

In total, 70 college students were recruited in two countries. Thirty-four students (male = 27; female = 7) were recruited using the online recruitment system at MTU. Thirty six students (male = 15; female = 21) were recruited using the online recruitment system at Korean German Institute of Technology. Auditory stimuli were presented via computers and headphones (Sennheiser HD 380 Pro headset). The auditory stimuli were composed of two categories: 1) auditory icons and 2) earcons. There was no visual stimulus in the experiment. Each participant listened to all sound clips from both of the sound categories. Sound clips for each affective adjective range from two to seven (M = 3.73 for auditory icons, M = 3.83 for earcons). They could listen to the same sound repeatedly as much as they wanted. After listening to all sounds, participants were asked to record which of the sound clips best conveyed a specific affective adjective (e.g., angry, boring, etc.). Upon completion of the task for one category (e.g., auditory icons), participants did the same for the other category (e.g., earcons). The order of category (auditory icons and earcons), affective adjectives, and sound clip presentation were randomized. Finally, participants were asked to decide which better conveyed the specific emotion between their favorite in auditory icons and favorite in earcons. Each session lasted around 60 minutes.

4. RESULTS AND DISCUSSION

Overall, in the same category (either auditory icons or earcons), participants in both countries showed similar preference for sounds. In other words, similar sounds induced similar emotions for a majority of both participants. In auditory icons, participants selected the same sound as their favorite for 23 emotions out of 30. In addition, three sounds selected as Americans' favorites were selected for Koreans' second best with the similar number of participants. Taken together, 26 (87%) emotional sounds were similarly selected as their favorite in both countries. They selected different sounds only for the remaining 4 emotions: boring, calm, confused, and depressed. For earcons, they preferred the exactly same sounds for 25 emotions (83%). They selected different sounds for the remaining 5 emotions: calm, free, impressive, surprising, and uneasy.

There were different trends in terms of the best selection across sound categories. Koreans showed stronger preference for either auditory icons or earcons, whereas Americans showed more distributed preference between the two categories. To illustrate, to express 'angry', 92% Korean participants preferred auditory icons (traffic jam) and only 8% preferred earcons (distorted guitar). However, 52% American participants preferred auditory icons and 48% preferred earcons. In fact, nine auditory icons were significantly more preferred than earcons by Koreans, but three auditory icons were significantly preferred by Americans (determined by chi-square goodness of fit tests, $p < 0.05$). Likewise, eight earcons were significantly more preferred than auditory icons by Koreans, but five earcons were significantly preferred by Americans (see Tables 1&2). For 'fresh', 'lively', and 'pleasant (happy)', Americans and Koreans commonly preferred auditory icons over earcons. For 'fresh', 'water pouring into an ice-filled glass' sound was commonly selected as the best. For 'lively', 'cheering and applauding crowd' sound was commonly selected as the best. For 'pleasant', 'child laughing' sound was commonly selected as the best. For 'dreamy' and 'fancy', Americans and Koreans commonly preferred earcons over auditory icons. For 'dreamy', 'whole tone scale' sound was commonly selected as the best. For 'fancy', 'baroque style harpsichord' sound was commonly selected as the best. We might infer that we could utilize these sounds as standardized auditory emoticons for both countries. However, other than these five sounds, there was no commonality in categorical preference.

There were not clear results regarding basic emotions, but we found some trends. Basic emotions tend to be more mapped onto auditory icons than earcons. For 'happy' (child laughing) and 'disgusting' (man vomiting), both Americans and Koreans tended to prefer auditory icons. For 'angry' (traffic jam) and 'surprising' (man short gasp), Koreans preferred auditory icons. However, if there is a typical association between an affective state and a musical parameter, earcons were preferred. For 'depressed (sad)', both Americans and Koreans tended to prefer earcons (minor chord). For 'fearful', Americans preferred earcons (tremolo string sound). This might be because both countries have a very similar structure of western music education from the elementary school. We might explain some of the results based on cultural differences. For example, Korean participants (who live in Seoul where the traffic condition is the worst in the world) always suffer from traffic jam and could more clearly associate traffic sound with their angry memory. On the other hand, American participants (who are exposed to specific media, e.g., Hitchcock's movie "Psycho") could more easily associate the tremolo string sound with the fearful emotion.

Table 2. Participants' preference between the two auditory cue types.

	USA	South Korea
Auditory Icons	Fresh Lively Pleasant	Angry Boring Disgusting Dynamic Fresh Lively Pleasant Surprising Uneasy
Earcons	Dreamy Fancy Intimate Modern Scared	Comfortable Depressed Dreamy Fancy Free Impressive Magnificent Soft

5. CONCLUSION AND FUTURE WORKS

We hypothesized that we would identify commonalities and differences in the relationship between affective sounds and affective keywords in the two different countries. In a single sound category, people mostly (over 80%) chose the same sound for a certain affective state. In other words, people prefer the same sound in a given sound set, regardless of culture or country. However, if they are given two different sets of sounds (i.e., auditory icons and earcons), their preference selection might vary. While Americans' preference was distributed more across the two categories, Koreans' preference was converged more towards either auditory cue type. We cautiously infer that it might reflect the cultural differences between the two countries, in terms of individualism-collectivism. However, of course, there are more variables that we cannot be sure for control, including the effects of knowledge, experience, memory, media, and other individual differences. Note that because all of our sounds were created by Americans, there might be inherent biases in the sound sets. Simply asking users about their preference may not be the best way to design a good user interface, but we believe that we could learn, at least, how to avoid the worst sound for both countries. This effort is expected to have substantial implications for designing culture-specific auditory user interfaces as well as standardized auditory emoticons, or the sounds per se that work for all. Designers and researchers could get enough hints for the next design iteration from this initial result.

Despite the successful initial attempt, there were some limitations inherent in the current study. The number of participants in each country was quite small and thus, may not be sufficient to draw a firm conclusion. We are still under data collection and the present paper is a type of summary of initial findings. Moreover, the participant groups were mostly graduate or undergraduate students and thus, they might not be representatives of all the populations. However, this homogeneous population can also serve as a controlled variable. Another cultural limitation may be that participants' importance of emotions and the meaning of affective adjectives may not be the same across the two groups.

Our next step includes recruiting larger samples with multiple generations and more balanced gender in each country and hopefully, extending our project team to more diverse cultures and countries, including Europe, Africa, etc. In addition to the lab study, we will construct an auditory-specific affect dimension as 2D or 3D coordinates based on our preference data [c.f., 13, 14] and implement the interactive version online so that users from any cultures or countries can contribute to adding any type of emotional sounds where appropriate in a given coordinate. Of course, they will be allowed to listen to the current sounds and freely download them for their own research or design.

6. REFERENCES

[1] G. Kramer, *An introduction to auditory display.* In G. Kramer (Ed.), Auditory display: Sonificaiton, audification, and auditory interfaces, MA, USA: Addison-Wesley, 1994.

[2] W. W. Gaver, "Auditory icons: Using sound in computer interfaces," *Human-Computer Interaction,* vol. 2, pp. 167-177, 1986.

[3] M. M. Blattner, D. A. Sumikawa, and R. M. Greenberg, "Earcons and icons: Their structure and common design principles," *Human-Computer Interaction,* vol. 4, pp. 11-44, 1989.

[4] W. W. Gaver, "The SonicFinder, a prototype interface that uses auditory icons," *Human-Computer Interaction,* vol. 4, pp. 67-94, 1989.

[5] S. A. Brewster, "The design of sonically-enhanced widgets," *Interacting with Computers,* vol. 11, no. 2, pp. 211-235, 1998.

[6] B. N. Walker, J. Lindsay, A. Nance, Y. Nakano, D. K. Palladino, T. Dingler, and M. Jeon, "Spearcons (speech-based earcons) improve navigation performance in advanced auditory menus," *Human Factors,* vol. 55, no. 1, pp. 157-182, 2012.

[7] M. Jeon and B. N. Walker, ""Spindex" (Speech Index) improves acceptance and performance in auditory menu navigation for visually impaired and sighted users," *ACM Transactions on Accessible Computing*, vol. 3, no. 3, pp. 10:11-26, 2011.

[8] R. Schleicher, S. Sundaram, and J. Seebode, "Assessing audio clips on affective and semantic level to improve general applicabiility," in *Proc. of DAGA*, 2010.

[9] P. M. C. Lemmens, A. De Haan, G. P. van Galen, and R. G. J. Meulenbroek, "Emotionally charged earcons reveal affective congruency effects." *Ergonomics,* vol. 50, no. 12, pp. 2017-2025, 2007.

[10] P. Larsson, A. Opperud, K. Fredriksson, and D. Västfjäll, "Emotional and behavioural response to auditory icons and earcons in driver-vehicle interfaces," in *Proc. of the Int. Technical Conf. on the Enhanced Safety of Vehicles* 2009.

[11] P. M. C. Lemmens, "Using the major and minor mode to create affectively-charged earcons," in *Proc. of the Int. Conf. (ICAD2005),* 2005.

[12] A. Pirhonen, K. Tuuri, M.-S. Mustonen, and E. Murphy, "*Beyond clicks and beeps: In pursuit of an effective sound design methodology,"* In I. Oakley and S. Brewster (Eds.), *HAID 2007. LNCS*, vol. 4813, pp. 133-144, Berlin Heidelberg, Germany: Springer-Verlag, 2007.

[13] J.-H. Lee, M. Jeon, and K. H. Han, "The analysis of sound attributes on sensibility dimensions," in *Proc. of the Int. Conf. (ICA2004),* 2004.

[14] M. Jeon, J.-H. Lee, Y. E. Kim, and K. H. Han, "Analysis of musical features and affective words for affection-based music search system," in *Proc. of the Conf. (KCCS2004),* Seoul, Korea, 2004.

[15] P. Ekman, "An argument for basic emotions," *Cognition and Emotion*, vol. 6, pp. 169-200, 1992.

[16] M. McGee-Lennon, M. K. Wolters, R. McLachlan, S. Brewster, and C. Hall, "Name that tune: Musicons as reminders in the home," in *Proc. of the Int. Conf. (CHI11)*, 2011.

[17] M. Jeon, "Lyricons (Lyrics + Earcons): Designing a new auditory cue combining speech and sounds," In C. Stephanidis (Ed.), *Posters, Part I, HCII 2013, CCIS,* vol. 373, pp. 342-346, 2013.

Table 1. The descriptions of the most preferred sounds in each category and percentage of the selection between the two categories. * indicates p-values < 0.05.

Description of Most Preferred Auditory Icons	Auditory Icons Preferred (USA)	Auditory Icons Preferred (Korea)	Affective Adjectives	Description of Most Preferred Earcons	Earcons Preferred (USA)	Earcons Preferred (Korea)
Traffic Jam	52%	**92%***	**Angry**	Distorted percussive guitar chords	48%	8%
Sigh (Yawning)	55%	**73%***	**Boring**	Descending base	45%	27%
Breeze through trees and birds chirping (Bell)	52%	32%	**Calm**	**Dreamy pad (Clarinet sound)**	48%	68%
Wind and shivering	67%	46%	**Cold**	Wind and descending piano notes	33%	54%
Sigh of relief and creaking of chair as sinking in	61%	27%	**Comfortable**	Woodwind chords	39%	**73%***
Quizzical grunt (Magic sounds)	55%	46%	**Confused**	Pitch bent tuning fork	45%	54%
Thunder clap or Distant ominous sound	58%	54%	**Dark**	Ominous descending strings	42%	46%
Glass breaking	45%	65%	**Delicate**	High-pitched oscillating piano notes	55%	35%
Dog whimpering (Male weeping)	39%	27%	**Depressed (sad)**	Piano minor chords	61%	**73%***
Man vomiting	64%	**97%***	**Disgusting**	Descending deep synthesized tones	36%	3%
Synthetic pulsing	6%	27%	**Dreamy**	Whole tone scale	**94%***	**73%***
Crowd cheering	39%	**76%***	**Dynamic**	2 high pitched trumpet sounds	61%	24%
Spoon tapping champaign glass	30%	0%	**Fancy**	Baroque style harpsichord	**70%***	**100%***
Wings flapping and bird chirping	64%	0%	**Free**	**Synthesized choir and chime (Ascending flute)**	36%	**100%***
Water pouring into an ice-filled glass	**70% ***	**97%***	**Fresh**	Funk music baseline	30%	3%
Grating metal	42%	57%	**Harsh**	Combination of high pitched keyboard notes	58%	43%
Amazed "woah"	55%	0%	**Impressive**	**Trumpet fanfare (Orchestra ending and symbals)**	45%	**100%***
Girl pleased "ooh" or Kissing	18%	69%	**Intimate**	Aura (pad) and bass plus snare	**82%***	31%
Cheering and applauding crowd	**70% ***	**81%***	**Lively**	Ascending synthetic violin with percussion	30%	19%
Trumpet fanfare or Thunder clap	45%	0%	**Magnificent**	Synthesized choir	55%	**100%***
Typing, and cacophony of beeping	24%	35%	**Modern**	Fuzzy pad and staccato melody	**76%***	65%
Typing on keyboard	36%	62%	**Plain**	Single flute note	64%	38%
Child laughing	**70%***	**97%***	**Pleasant (happy)**	3 ascending piano notes	30%	3%
Woman blood curdling screaming	30%	46%	**Scared (fearful)**	Tremolo string sound	**70%***	54%
Single tick of clock	48%	41%	**Simple**	Rhythimic xylophone	52%	59%
Wobbly bell	42%	19%	**Soft**	Descending piano (with reverb)	58%	**81%***
Loud bang	42%	54%	**Strong**	Synthetic bass drum	58%	46%
Man short gasp	52%	**81%***	**Surprising**	**Ascending fuzzy keyboard (Orchestration bang)**	48%	19%
Scraping fingernails on chalkboard or Screechy sound	36%	**73%***	**Uneasy**	**Tremolo keyboard (Trombone sound)**	64%	27%
Fire crackling	67%	54%	**Warm**	Acoustic guitar chords	33%	46%